CONCEPTUAL DATA MODEL DOCUMENT (**DRAFT**)
November 21, 2003

Purpose
The purpose of this model is to have a complete and concise picture of **all** of the BLM's spatial data. It is not a database model, but rather an organizational model which provides a logical structure for existing data and, just as important, for new data being created every day. Using the model will allow the spatial database to grow in an organized fashion and control complexity and duplication. The goal is to migrate toward simpler data organization which is easier to understand, use and maintain.

Intended Audience
This document is intended to give managers and data stewards information about the OR/WA Data Design project, in particular the progress and status of the work on the conceptual data model. It is also intended to provide a basic documentation of the work that has been accomplished. It is envisioned that this document will be updated as additional steps are taken. It will become a final document once work on the conceptual model is completed and the effort moves on to development of the logical data model.

Methodology
Data Design Team – As part of the ArcGIS Transition Plan there was a sub-task identified for "Data Standards, Geodatabase Design, Metadata". The goals for this team are to: 1) migrate data from the existing ArcInfo Workstation library structure into the ArcGIS geodatabase structure and 2) redesign our data to take advantage of the relational database nature of the geodatabase. In November 2001 a team was established to address this task. Since that time there have been additions to the original team as the activities moved from the initial inventory stage, to rehosting of data in SDE, and finally to initial conceptual data model. Current team members are: Stan Frazier, Dan Wickwire, Georgia Bosse, Arthur Miller (all from OR-955), Pam Keller (Burns District), Debbie Smith, Barbara Welsh, and Steve Salas (Titan Contractors), Alan Rhodes and Mark Koski (Salem District), Duane Dippon, Janis VanWyhe, Jim Alegria, and Chris Cadwell (Division of Renewable Resources, Marc Thomas (Division of Management Services), and Kate Taylor (ESRI).

Initial Data Groups/Categories – The first design task that the team addressed was to identify data subject areas (high level groupings of data). An early decision point was whether to categorize data based on the use of the data or on the intrinsic qualities of the data. The initial set of data subject areas (October 2002) included 16 groups (e.g. archaeology, climate, fire, lands, etc) which are more a reflection of data use than intrinsic qualities. After giving this more thought it was decided that the highest level of data subject areas should be based on intrinsic qualities (does it occur naturally, is it a physical thing but built by humans, is it a human concept that may not have any physical appearance). The top level of the data subject areas were changed to just four data types:

a. **Resources** - Natural. Exist physically. Attributes include factual data;
b. **Facilities** - Human structures. Exist physically. Built on top of resources;
c. **Boundaries** - Human constructs. Often no physical existence;

d. **Activities -** Human. Attributes include target, benefit, costs.

Data groupings in the lower (2nd, 3rd, etc) levels are based on use similar to the initial data subject areas. This formed the basis for the first conceptual data model.

Conceptual Model Concepts – A team meeting was held August 27-28, 2003 where we discussed the data design project, what had already been done, developed a project plan, and discussed the next step to take. The development of an initial conceptual data model was that next step. Pam Keller was tasked to create this 1st cut at a conceptual model. In our team meeting the ESRI contractor stated that this model should go to the "entity" level and when asked what she meant by entity, she said "feature class" which is a collection of related spatial things (features) with the same geometry (a group of related point features or a group of related line features).

One of the main purposes of a conceptual model is entity definition.

Entities in a GIS data model are spatial and their definitions are dependent on their spatial characteristics.

The model attempted to capture **all** the **unique** spatial entities and in addition the model shows each unique spatial entity only **once**. Three examples:

a. Fish species could be listed under RESOURCES-SPECIES OCCURRENCES but fish are spatially "attached" to streams which are already listed under RESOURCES-WATER
b. Recreation use data could be associated with any number of boundaries (campgrounds, wilderness areas, resource areas, etc.), all of which are already accounted for in the model so recreation use data does not create a new entity.
c. Weed infestations are spatially defined differently from areas surveyed for weeds and there must be two distinct spatial entities even though there will be many duplicate areas between the two entities.

Attributes associated with spatial entities are stored in subject oriented tables (e.g. recreation information, fish species information, etc.) and linked to the spatial entities. There will be tables containing very different subject matter but linked to the same spatial entity. The model does **not** attempt to show these tables, only the spatial entities. The model does attempt to show **core** attributes that are directly attached to the spatial entity table. Core attributes include: linking fields, attributes commonly used by many user groups and attributes which lead to new features (lead to line splits or new polygons).

Geometry is one of the primary characteristics of spatial entities. In addition to the fundamental geometry of point, line, polygon or raster there is a 2nd level geometry that must be described. It defines whether the features are distinct and isolated or form a continuous, connected layer on the landscape and whether it is necessary to have overlapping features. Overlapping lines or polygons are

complex geometry and are only used when the spatial definition of the data involves non-finite overlap (as with tracking continuous change over time).

With all of the above as background, model groups were based on a number of considerations including:

Core attribute commonality
Geometry
Subject matter
Update characteristics
Query and analysis

There have been three iterations of the conceptual model so far. Additional iterations will occur as future reviews are completed.

Accessory Model - This model steps back from the more detailed conceptual model and depicts the data with a different view. The overall goal is the same: to have a complete and concise picture of **all** of the BLM's GIS data, but the organization is not based on any physical implementation. The conceptual model groups data entities according to commonalities of definition, spatial characteristics, core attribute and query requirements and data maintenance considerations. This accessory model groups data entities by subject matter commonality. It is therefore closely tied to BLM program areas such as Fire, Wildlife, Minerals. Each of the program areas will be found in all four of the top level categories (Resources, Facilities, Boundaries and Activities). This model was developed to aid in communications with data stewards and program managers. The four diagrams that make up the Accessory Model can be found in Appendix 2.

Presentations – Two presentations have been made to begin the dialogue with data stewards and GIS specialists. The first was a presentation at the Data Stewardship Workshop (Oct 1-2, 2003). This was a short presentation for about 20 data stewards to describe the intent of this work and to show some examples of the models. The 2nd draft model was presented at this workshop. The second presentation was at the OR/WA Field User Group meeting (Oct 14-16). This presentation was very similar to the one done for data stewards. The 3rd draft model was presented at this meeting. One of the major outcomes of this presentation was the need to develop the Accessory Model (described above).

Current Conceptual Data Model-Summary of Changes

The current conceptual model is revision 3 for all groups except cadastral boundaries where the current version is revision 3a (this was a minor adjustment to version 3 to add "special management areas-critical habitat" to the model). Major changes that have occurred between development of the initial conceptual model and the current model are:

- Established primary model concepts.
- Defined the meaning of entity. Model to capture all unique spatial entities, show each unique entity only once.
- Show only necessary core attributes for each entity, not the full scope of information that might be used in relation to the entities.

- Recognize the nature of entity geometry. Is entity inherently continuous or consisting of isolated occurrences over areas of interest? Does entity overlap similar entities, etc.

- With this information explored, the base model groups were formed taking into account core attribute commonality, geometry, subject matter, update characteristics, query and overlay users.

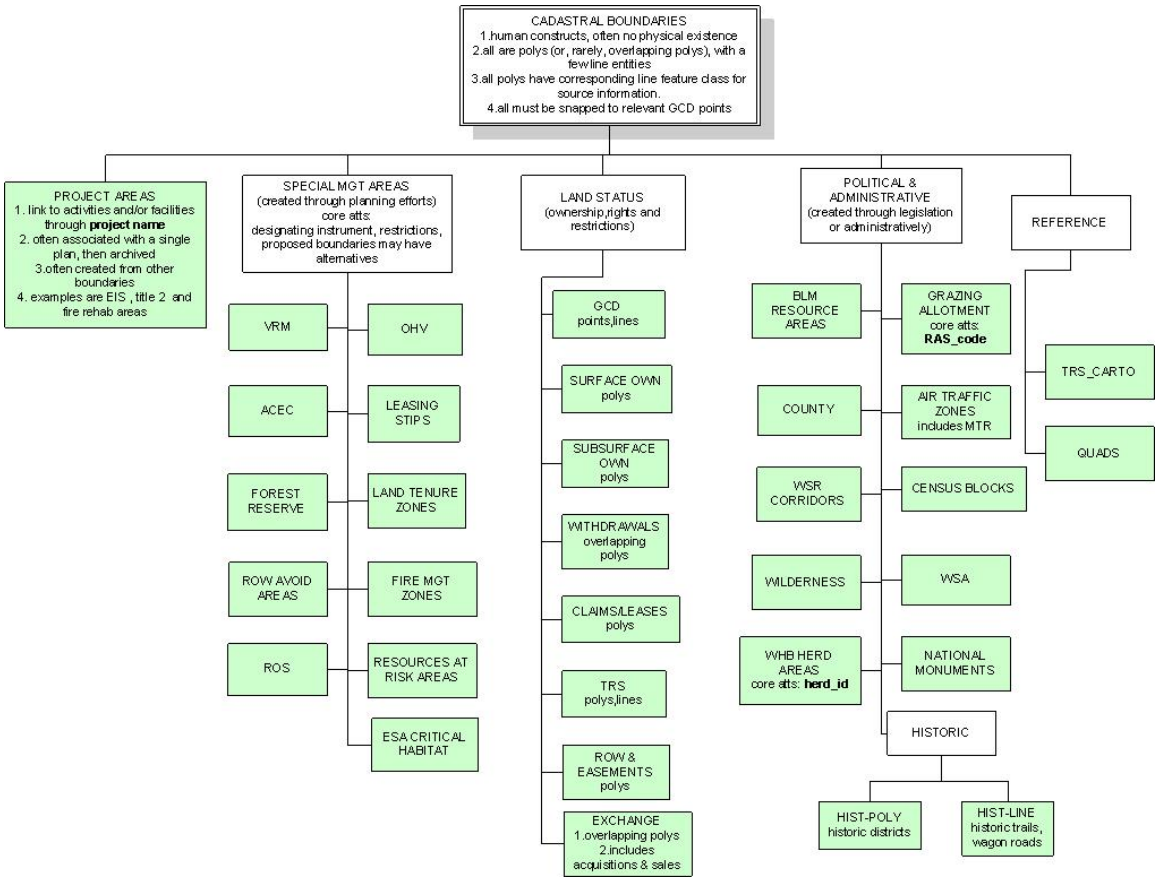
- Clarification of terminology. Many minor changes, most toward more precise terminology (ie. Boundaries changed to Cadastral Boundaries).

- Clarification of the spatial concept behind Activities entities. These entities represent location and area of where activities occur rather than data resulting from of activities.

The detailed information on changes by model version can be found in Appendix 3.

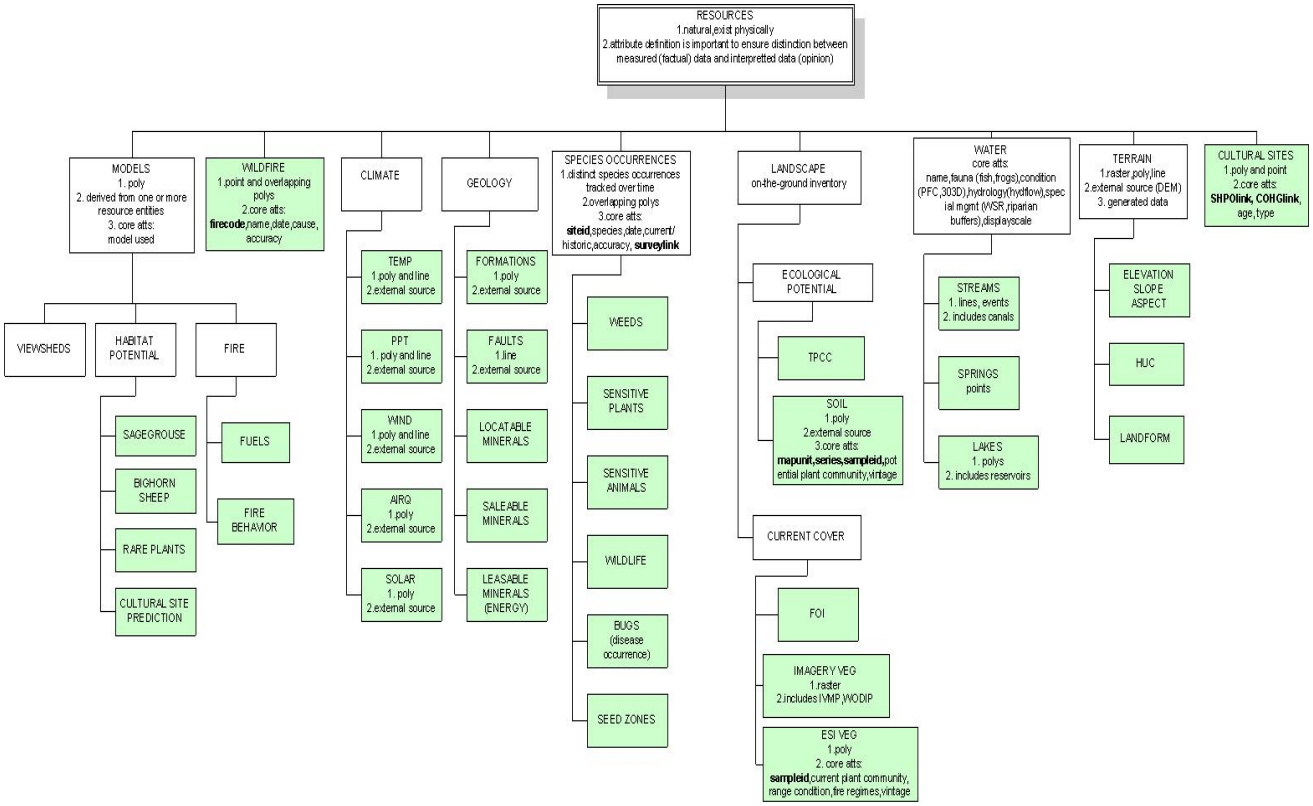
What follows are the current models for each Data Subject Area followed by definitions for each major entity in that model.

Figure 1 - Cadastral Boundaries



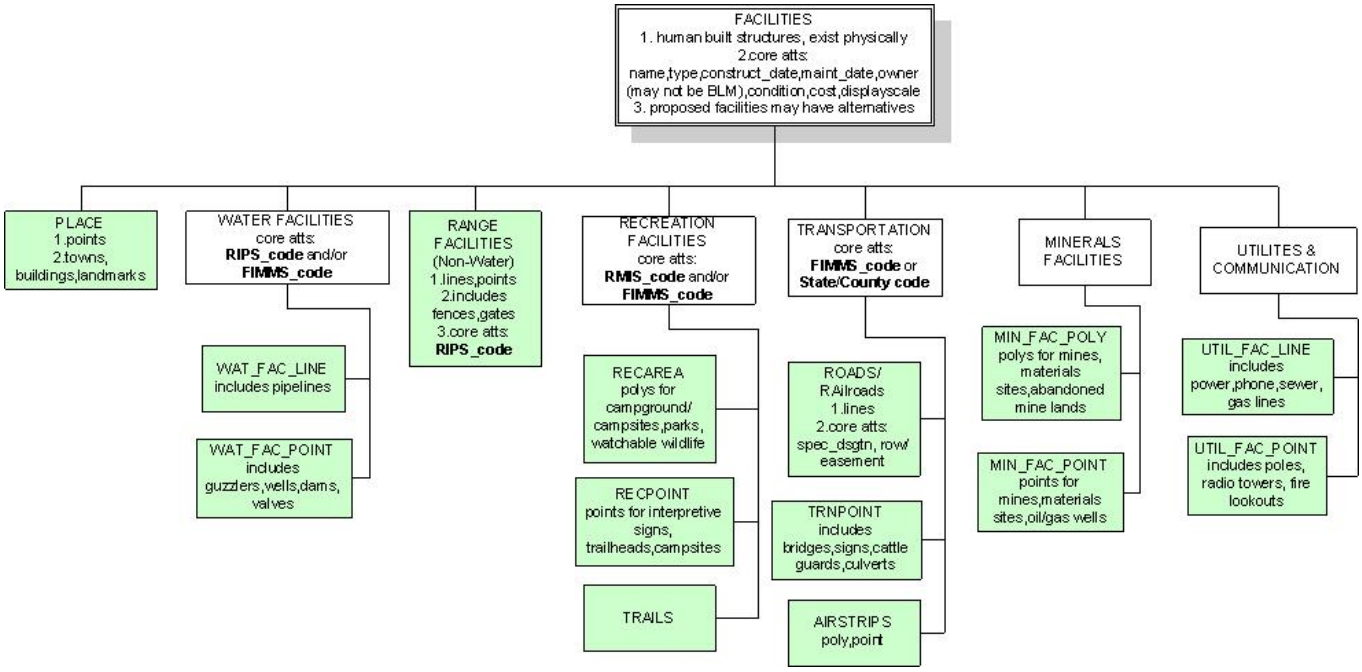
Special Management Areas –
Land Status –
Political and Administrative –
Reference –
Project Areas –

Figure 2 - Resources



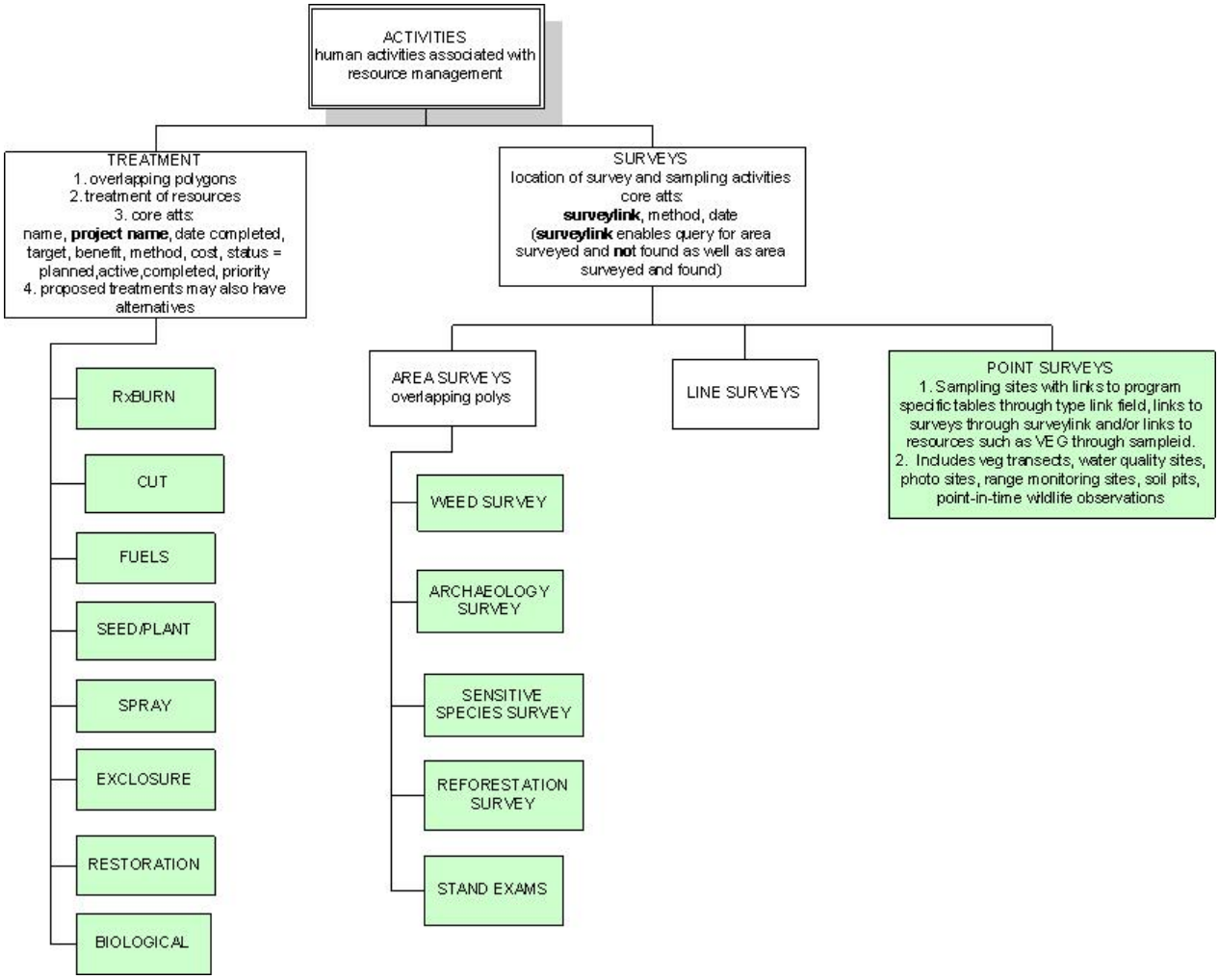
Models –
Wildfire –
Climate –
Geology –
Species Occurrences –
Landscape –
Water –
Terrain –
Cultural Sites –

Figure 3 - Facilities



Place –
Water Facilities –
Range Facilities –
Recreation Facilities –
Transportation –
Minerals Facilities –
Utilities and Communication –

Figure 4 - Actitivities



Treatment –
Surveys –

Next Phases

1. Data Steward Review – The next step in development of the conceptual data model is to get in-depth review by the data stewards. This is needed to make sure that the data needed by data stewards in their day-to-day work is accounted for as well as making sure definitions and relationships are correct.
2. Logical Data Model – Once the conceptual model is completed the next phase is development of the logical data model. This model contains a greater level of detail, especially with regards to relationships between entities and attributes of each entity. At this stage geometry types are defined, attributes, subtypes, spatial integrity rules, domains, cartographic reference rules, etc are defined.
3. Project Plan – The project plan is fluid and changes quite often (especially in the area of due dates). The current plan (at a fairly high level) is shown below.

	Task Name	Duration	Start	Finish
1	**Conceptual Design**	**140.5 days**	**Mon 9/1/03**	**Mon 3/15/04**
2	Initial Cut at Conceptual Model	5 days	Mon 9/1/03	Fri 9/5/03
3	Review by Data Team	5 days	Mon 9/8/03	Fri 9/12/03
4	ID Data Team	5 days	Tue 9/2/03	Mon 9/8/03
5	**Kick Off Meeting**	**3 days**	**Mon 9/15/03**	**Wed 9/17/03**
9	Create Conceptual Model Document	16 days	Fri 10/31/03	Fri 11/21/03
10	Review by Data Team	4 days	Mon 11/24/03	Thu 11/27/03
11	Present to Data Stewards	2 days	Fri 11/28/03	Mon 12/1/03
12	**Develop Conceptual Model**	**44 days**	**Tue 12/2/03**	**Fri 1/30/04**
23	Data Steward Introductory Meeting	0.5 days	Mon 2/2/04	Mon 2/2/04
24	Data Steward Interviews	30 days	Mon 2/2/04	Mon 3/15/04
25	Data Team Review Meeting	1 day	Tue 1/13/04	Tue 1/13/04
26	Update Model based on Comments	10 days	Wed 1/14/04	Tue 1/27/04
27	Presentation at Winter FUG Meeting	3 days	Tue 2/17/04	Thu 2/19/04
28	Presentation to Management	0.5 days	Tue 3/2/04	Tue 3/2/04
29	Notice to Proceed	0 days	Tue 3/2/04	Tue 3/2/04
30	**Logical Design**	**84 days?**	**Fri 2/20/04**	**Wed 6/16/04**
31	**Develop Logical Model - for each entity:**	**84 days?**	**Fri 2/20/04**	**Wed 6/16/04**
32	**Logical Design Methodology Workshop**	**1 day?**	**Fri 2/20/04**	**Fri 2/20/04**
38	**Develop Logical Models for Priority Layers**	**40 days**	**Mon 2/23/04**	**Fri 4/16/04**
49	Develop UML Model	40 days	Mon 3/22/04	Fri 5/14/04
50	Generate Schema	1 day	Mon 5/17/04	Mon 5/17/04
51	Define Metadata	5 days	Tue 5/18/04	Mon 5/24/04
52	Load Sample Data	5 days	Tue 5/25/04	Mon 5/31/04
53	**Schema Presentation Workshop**	**1 day**	**Tue 6/1/04**	**Tue 6/1/04**
57	Review & Comment Submission by Data Team	5 days	Wed 6/2/04	Tue 6/8/04
58	Update Model based on Review	5 days	Wed 6/9/04	Tue 6/15/04
59	Review by Stakeholders at June FUG	1 day?	Wed 6/16/04	Wed 6/16/04

APPENDIX 1 - DEFINITIONS

Definition of Terms Used

ArcGIS – An integrated scalable family of GIS software products from ESRI for data creation, management, integration, and analysis. Consists of (ArcCatalog, ArcMap, ArcToolbox, and more).

ArcInfo Workstation- Classic command line interface GIS from ESRI consisting of (ARC, ARCEDIT, ARCPLOT, ARC Macro Language [AML], and more).

SDE- Spatial Database Engine, software from ESRI that facilitates management of spatial data in a relational database management system.

Conceptual Data Model – Organizing principles that translate into functional descriptions of how phenomena are represented and related to one another.

Entity – A person, place, thing, or concept that we want to keep information about.

ESRI – Environmental Systems Research Incorporated is the company that makes the GIS software that the Bureau uses.

Feature – A specific instance (point, line, polygon) of an entity. A representation of a real world object.

Feature Class – A collection of features with the same type of geometry and the same type of attributes. When referring to geographic features, feature classes include point, line, area, and annotation.

Feature Dataset – A collection of feature classes that share the same spatial reference.

Geodatabase – (short for geographic database) is a physical store of geographic information inside a database management system (DBMS).

GIS –An information system that is designed to work with data referenced by spatial or geographic coordinates. ... both a database system with specific capabilities for spatially-reference data, as well [as] a set of operations for working with data. *From Jeffrey Star and John Estes, in Geographic Information Systems: An Introduction (Englewood Cliffs, NJ: Prentice-Hall, 1990), page 2-3:*

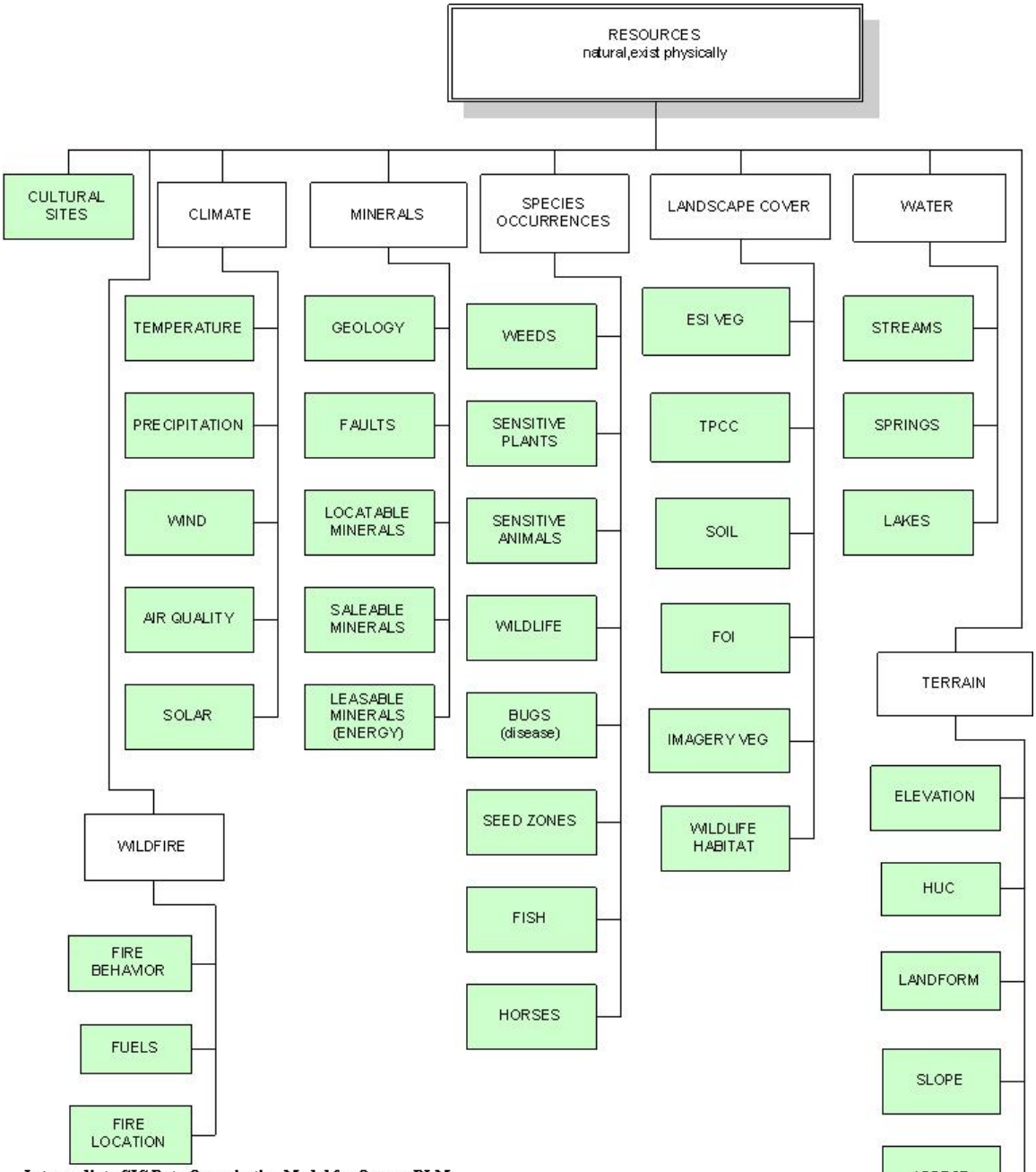
Logical Data Model – describes the data requirements to support a functional activity. An abstract representation of the structural data requirements of part or all of the operations of an organization.

Physical Data Model – describes an automated solution based on the physical environment and performance concerns. A framework for a physical database which may include: referential integrity, indexes ,views , alternate keys and other constraints, tablespaces and physical storage objects.

Relationship – An association or link between two objects in a database. Relationships can exist between spatial objects (features in feature classes), non-spatial objects (rows in a table), or between spatial and non-spatial objects.

Spatial – Data having an attribute of location in a definable coordinate system.

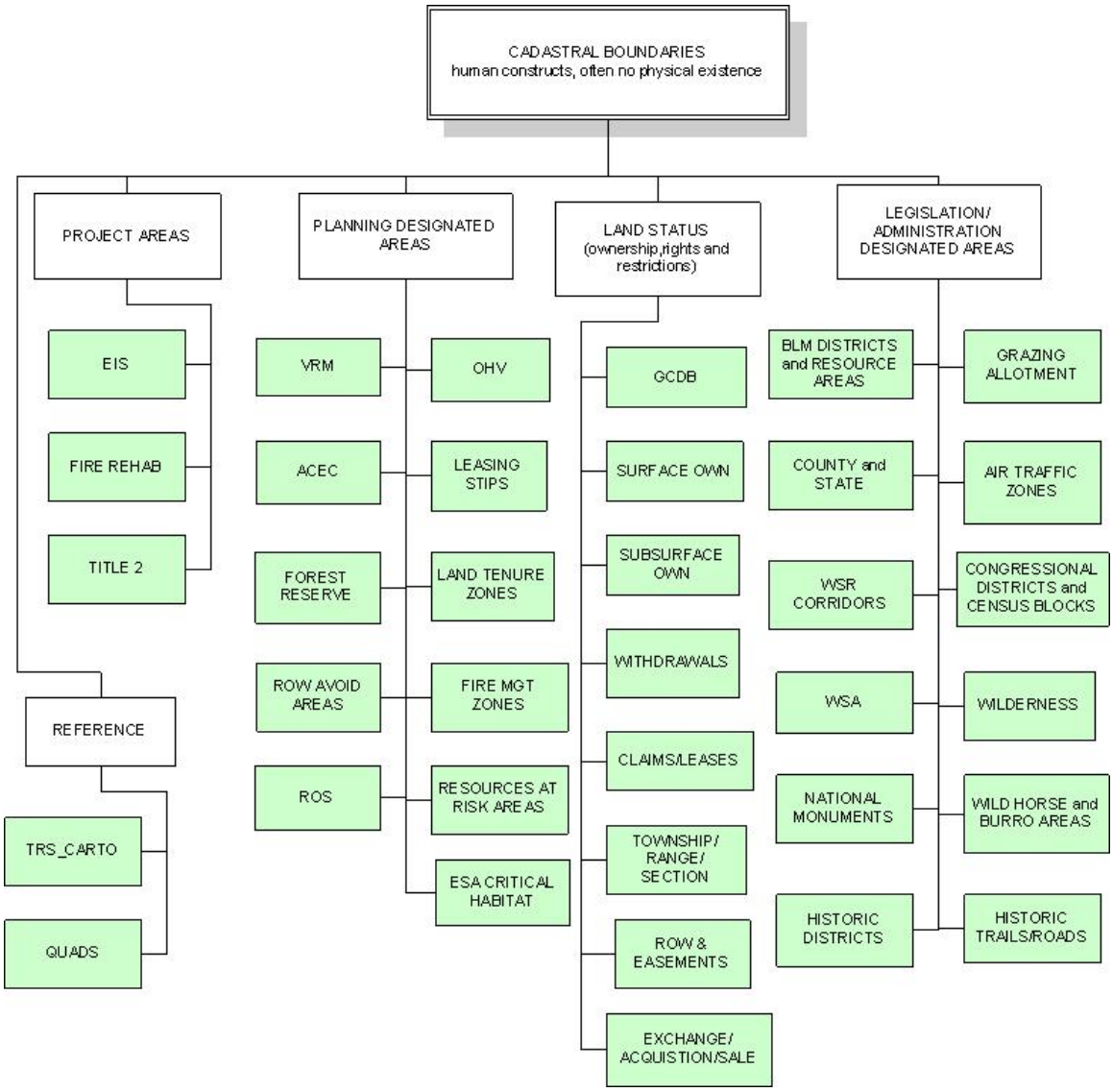
APPENDIX 2 – ACCESSORY MODELS



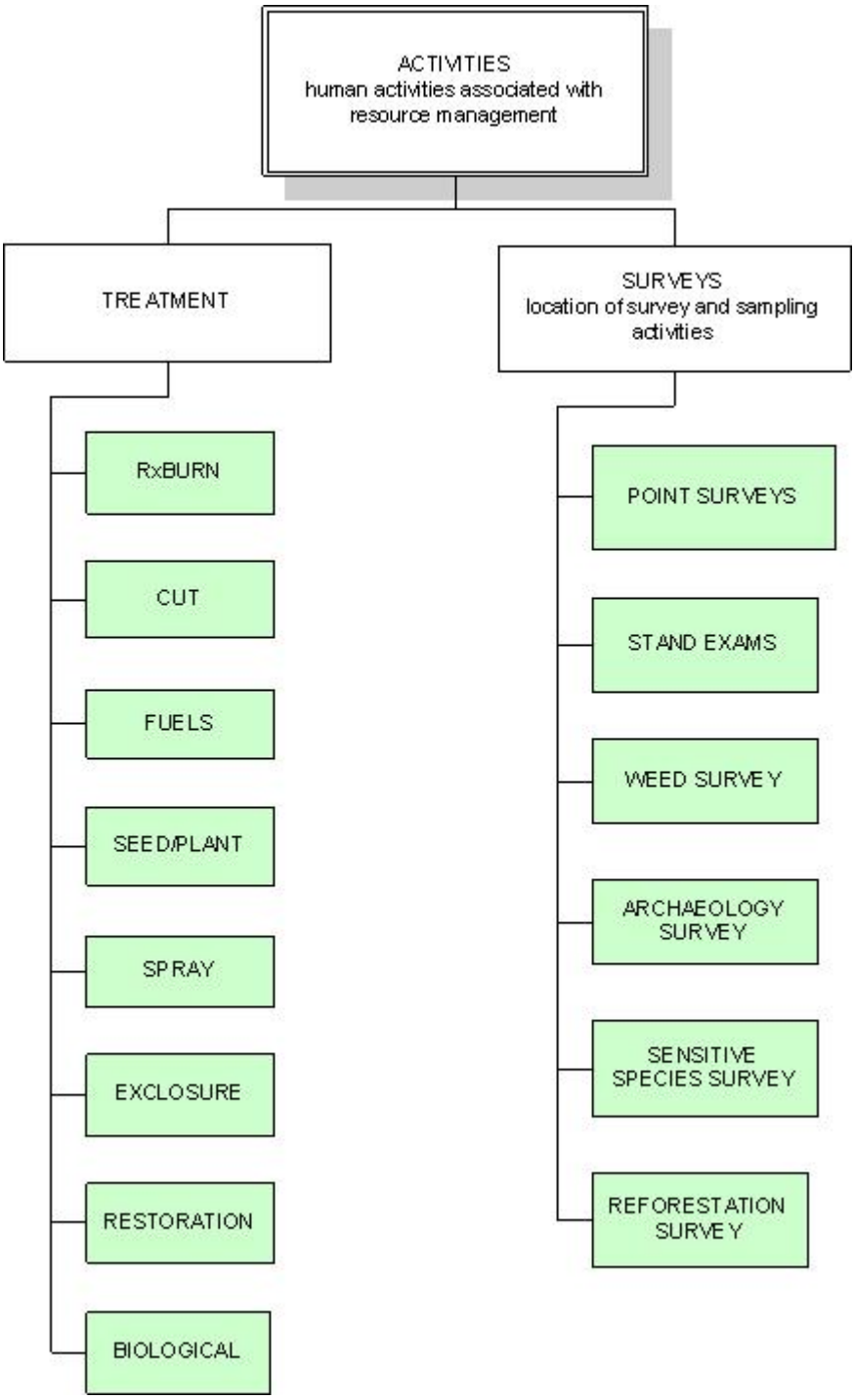
Intermediate GIS Data Organization Model for Oregon BLM
October 27, 2003

This model steps back from the more detailed logical entity model (September 2003). The overall goal is the same: to have a complete and concise picture of **all** of the BLMs GIS data, but the organization is not based on any physical implementation. The more detailed model groups data entities according to commonalities of definition, spatial characteristics, core attribute and query requirements and data maintenance considerations. This model groups data entities by subject matter commonality. It is therefore closely tied to BLM program areas such as Fire, Wildlife, Minerals. Each of the program areas will be found in all four of the top level categories (Resources, Facilities, Boundaries and Activities).

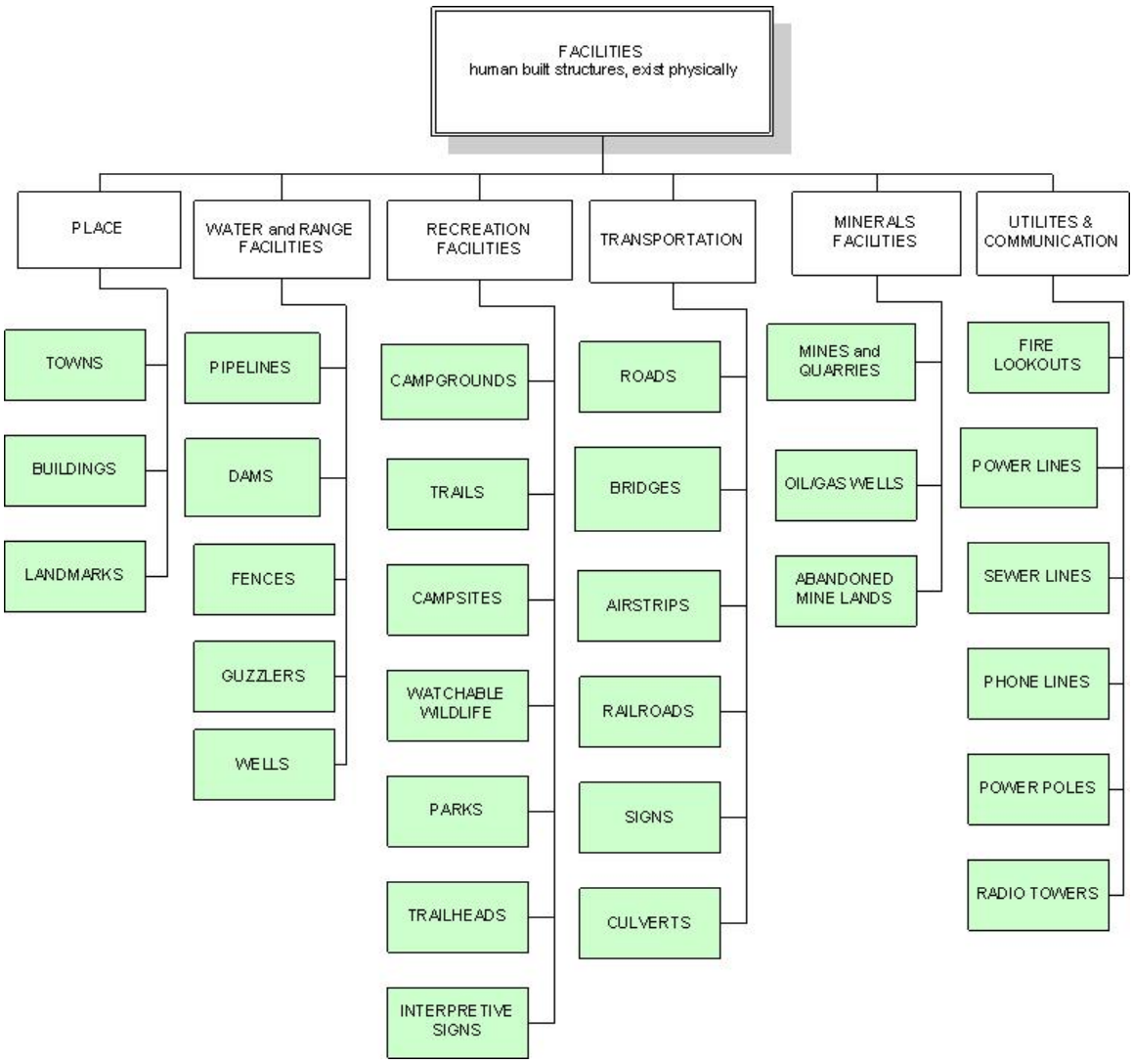
APPENDIX 2 – ACCESSORY MODELS



APPENDIX 2 – ACCESSORY MODELS



APPENDIX 2 – ACCESSORY MODELS



APPENDIX 3 – MODEL CHANGES DETAIL

Conceptual Model (2nd Draft) – Changes were made to the draft model following team review of the 1st Draft.

Under **Resources**:
SOLAR added to CLIMATE
ARCHAEOLOGY changed to CULTURAL
WATER streams clarified to include canals and lakes to include reservoirs
POPULATIONS changed to SPECIES OCCURRENCES, TIMBER STANDS removed and SEED ZONES added
Much discussion over LANDSCAPE COVER, existing vs potential vs modeled habitat. I attempted another view by putting 3 subgroups under LANDSCAPE for ECOLOGICAL POTENTIAL (soil and PNC veg), CURRENT COVER and MODELS (habitat, fuels, tpcc)
IMAGERY VEG added under LANDSCAPE , CURRENT COVER.

Under **Facilities**:
RANGE split out from WATER
Word "Facilities" added to subgroup titles (e.g. RECREATION FACILITIES) to avoid confusion.
CULTURAL (towns, landmarks) changed to PLACE

Under **Boundaries**:
Name Changed to **Cadastral Boundaries**
Definition of SPECIAL MGT AREAS clarified as "created through planning efforts" and POLITICAL & ADMINISTRATIVE as "created through legislation or administratively" and WSA, WSR Corridors, WILDERNESS, WHB Areas, HISTORIC districts and trails, NATIONAL MONUMENTS moved from the first to the second group
RESOURCES AT RISK areas added to SPECIAL MGT AREAS
FIRE MGT ZONES moved from POLITICAL to SPECIAL MGT
LEASES added to CLAIMS under LAND STATUS
New 2nd level categories added fro PROJECT AREAS and for REFERENCE

Under **Activities:**
Much discussion over ideas of "Survey", "Sampling" and "Monitoring"
MONITOR changed to SAMPLING LOCATIONS (Point) and
SURVEY changed to SURVEY AREA with link field added that links to species occurrence entities to answer the "looked and didn't find" question.

Conceptual Model (3rd Draft) – Additional changes were made after a team conference call to review the 2nd Draft.

Under **Resources:**
MODELS taken out from under LANDSCAPE and made a 2nd level category

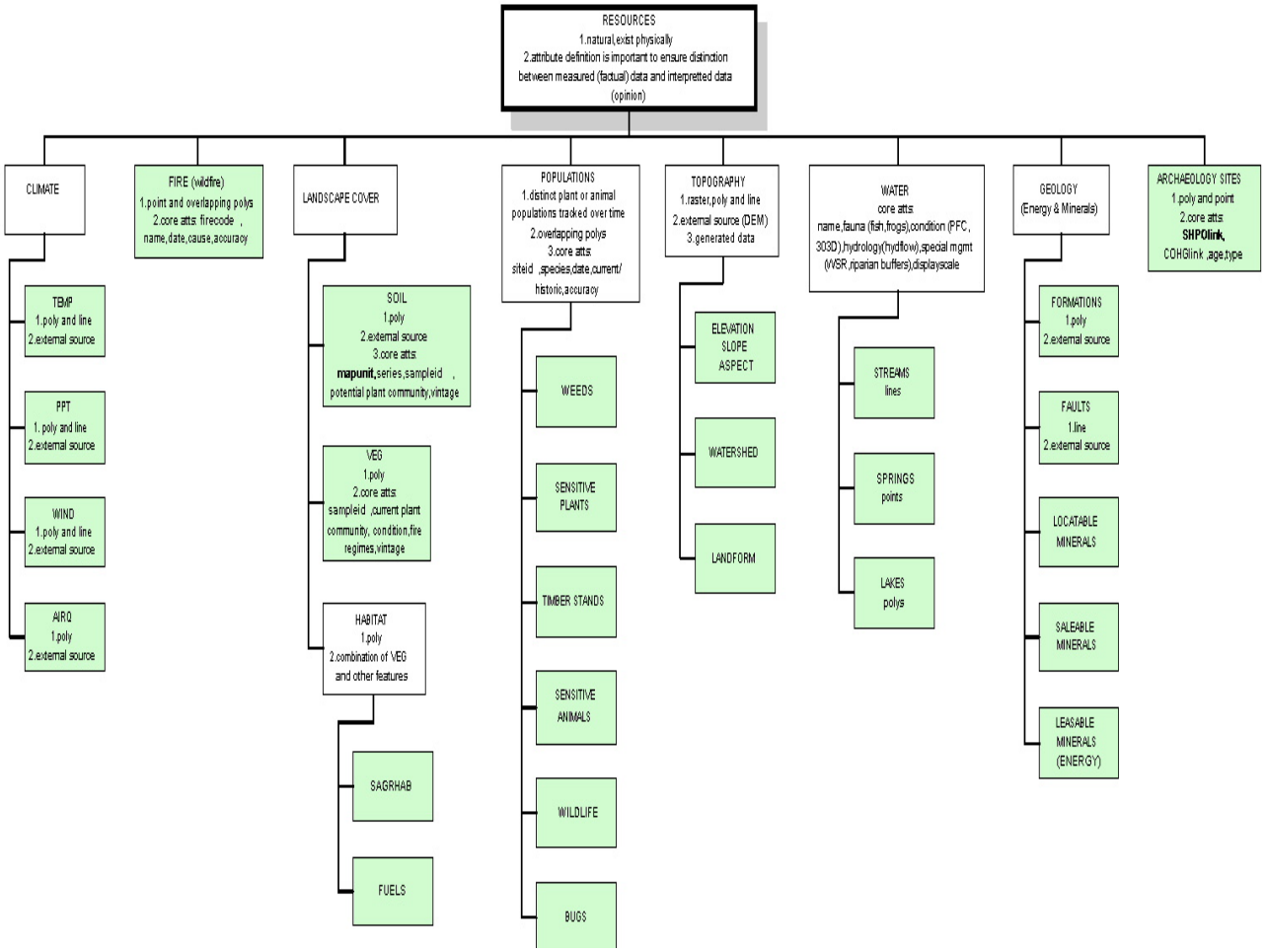
with subcategories for HABITAT POTENTIAL, FIRE and VIEWSHEDS
TPCC added under LANDSCAPE ECOLOGICAL POTENTIAL and PNC VEG put back into SOIL
FOI added under LANDSCAPE, CURRENT COVER.

Under **Facilities:**
AIRSTRIPS added under TRANSPORTATION

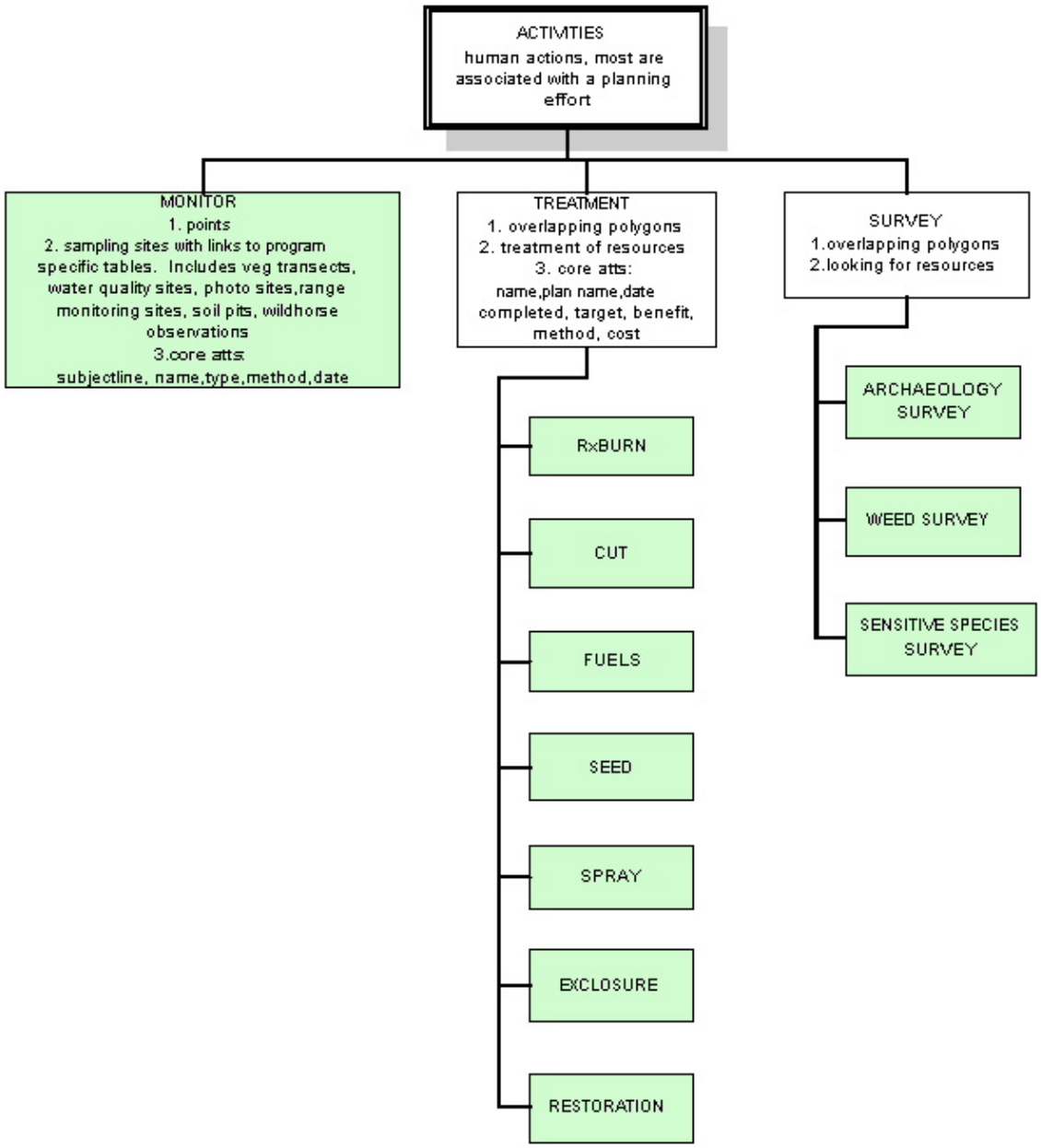
Under **Boundaries:**
PROJECT AREAS definition clarified
TRS_CARTO and QUADS added under REFERENCE
ESA CRITICAL HABITAT added under SPECIAL MGT AREAS after discussion with some Botanists and Biologists

Under **Activities:**
SAMPLING LOCATIONS and SURVEY AREAS combined into SURVEYS with 3 subcategories for AREA SURVEYS, LINE SURVEYS and POINT SURVEYS
REFORESTATION SURVEY and STAND EXAMS added under AREA SURVEYS
FOI removed from TREATMENT (moved to LANDSCAPE under Resources)
BIOLOGICAL added under TREATMENT

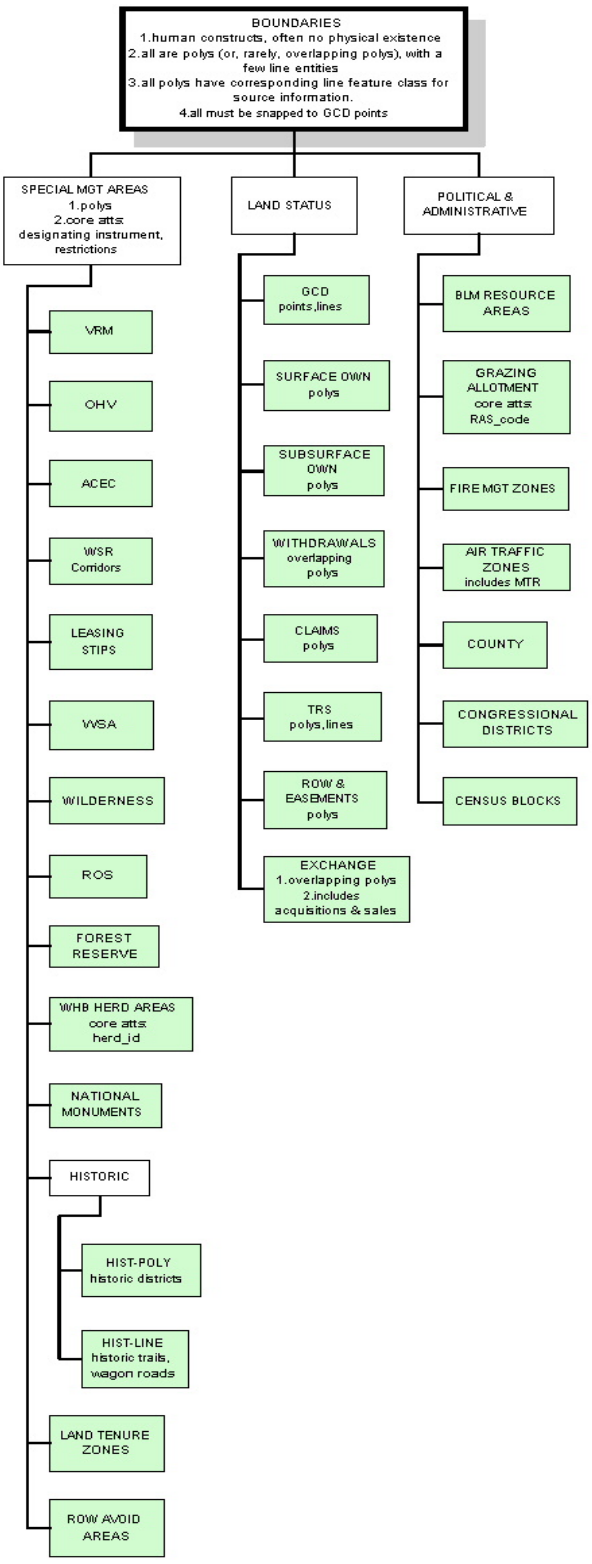
APPENDIX 4 - Initial Conceptual Models (version 1)



APPENDIX 4 - Initial Conceptual Models (version 1)



APPENDIX 4 - Initial Conceptual Models (version 1)


BOUNDARIES
1.human constructs, often no physical existence
2.all are polys (or, rarely, overlapping polys), with a few line entities
3.all polys have corresponding line feature class for source information.
4.all must be snapped to GCD points
SPECIAL MGT AREAS
1.polys
2.core atts: designating instrument, restrictions
LAND STATUS
POLITICAL & ADMINISTRATIVE
VRM
OHV
ACEC
WSR Corridors
LEASING STIPS
WSA
WILDERNESS
ROS
FOREST RESERVE
WHB HERD AREAS
core atts: herd_id
NATIONAL MONUMENTS
HISTORIC
HIST-POLY
historic districts
HIST-LINE
historic trails, wagon roads
LAND TENURE ZONES
ROW AVOID AREAS
GCD
points,lines
SURFACE OWN
polys
SUBSURFACE OWN
polys
WITHDRAWALS
overlapping polys
CLAIMS
polys
TRS
polys,lines
ROW & EASEMENTS
polys
EXCHANGE
1.overlapping polys
2.includes acquisitions & sales
BLM RESOURCE AREAS
GRAZING ALLOTMENT
core atts: RAS_code
FIRE MGT ZONES
AIR TRAFFIC ZONES
includes MTR
COUNTY
CONGRESSIONAL DISTRICTS
CENSUS BLOCKS

APPENDIX 4 - Initial Conceptual Models (version 1)